LAW OFFICES
LEYDIG, VOIT & MAYER, LTD.
TWO PRUDENTIAL PLAZA, SUITE 4900
CHICAGO, ILLINOIS 60601-6780

TELEPHONE: (312) 616-5600 TELECOPY: (312) 616-5700 (G3)
(312) 849-0495 (G4)

FACSIMILE COVER SHEET

DATE: NOVEMBER 18, 2004 NUMBER OF PAGES (INCLUDING THIS TRANSMITTAL COVER SHEET): 14

TIME:

YOUR REFERENCE: OUR REFERENCE: 211163

TO: Examiner Meltin Bell **(Courtesy Copy)**
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

TELEPHONE NUMBER: 571-272-3680

FACSIMILE NUMBER: **(571) 273-3680**

FROM: PHILLIP M. PIPPENGER **DIRECT LINE:** 312-616-5653

MESSAGE: PLEASE SEE THE ATTACHED AMENDMENT.

A confirmation copy of the transmitted document will:
- ☒ Not be sent. This will be the only form of delivery of the transmitted document.
- ☐ Be sent via First Class/Air Mail.
- ☐ Be sent via Overnight Courier

PATENT
Attorney Docket No. 211163

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Alex Gammerman

Application No. 09/831,262

Filed: June 29, 2001

For: DATA CLASSIFICATION APPARATUS AND METHOD THEREOF

Group Art Unit: 2121

Examiner: BELL, MELTIN

CERTIFICATE OF MAILING
I hereby certify that this INTERVIEW SUMMARY (along with any documents referred to as being attached or enclosed) is being deposited with the United States Postal Service on the date shown below with sufficient postage as first class mail in an envelope addressed to: Mail Stop AF, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.
Date: 11/18/04 Susan Matz

INTERVIEW SUMMARY

Mail Stop AF
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Applicants thank the Office for the courtesy of an interview held on Thursday, November 18, 2004. The participants discussed a proposed amendment, now submitted herewith, to address the objections to and rejections of the claims, e.g. under section 101, 112 and 103 (in view of Mizuno and Tsubaka). The amendment to the specification and claims to set out the meaning of "iid" was added to the amendment at the suggestion of the Office to clarify things a bit. At that point it was agreed that the amendments in the draft and those discussed didn't/wouldn't add new matter and should place the claims in condition for allowance. Again, applicants thank the Examiner and his Supervisor for their handling of the case and the interview.

Respectfully submitted,

Phillip M. Pippenger, Reg. No. 46,055
LEYDIG, VOIT & MAYER, LTD.
Two Prudential Plaza, Suite 4900
180 North Stetson
Chicago, Illinois 60601-6780
(312) 616-5600 (telephone)
(312) 616-5700 (facsimile)

Date: November 18, 2004

PATENT
Attorney Docket No. 211163
Date: November 18, 2004

In re Application of: Alex Gammerman
Application No. 09/831,262
Filed: June 29, 2001
For: DATA CLASSIFICATION APPARATUS AND METHOD THEREOF

Mail Stop AF
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Transmitted herewith is a response to an office action in the subject application.

☐ Small entity status is claimed for this application under 37 CFR 1.27.

☒ Petition for an extension of time for the period noted below, as well as for any additional period necessary to render the present submission timely. Please charge Deposit Account No. 12-1216 for the appropriate petition fee.

☐ Other:

☒ Please charge Deposit Account No. 12-1216 in the total amount indicated below. A duplicate copy of this transmittal sheet is enclosed herewith.

					SMALL ENTITY		OTHER THAN A SMALL ENTITY	
TIME EXTENSION PETITION FEE			three-month		$ 0.00		$980.00	
	subtract time extension fee previously paid		none		($ 0.00)		($ 0.00)	
CLAIM FEE	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR	EXTRA CLAIMS PRESENT	RATE	ADD'L CLAIM FEE	RATE	ADD'L CLAIM FEE
TOTAL	6	MINUS	9	= 0	x 9=	$	x 18=	$
INDEPENDENT	4	MINUS	4	= 0	x 44=	$	x 88=	$
☐ FIRST PRESENTATION OF MULTIPLE CLAIM					+ 150=	$	+ 300=	$
TOTAL AMOUNT TO BE CHARGED TO DEPOSIT ACCOUNT					TOTAL	$	TOTAL	$980.00

☒ The Commissioner is hereby authorized to charge any deficiencies in the following fees associated with this communication or credit any overpayment to Deposit Account No. 12-1216.
- ☒ Any filing fees under 37 CFR 1.16 for the presentation of extra claims.
- ☒ Any patent application processing fees under 37 CFR 1.17.

Respectfully submitted,
LEYDIG, VOIT & MAYER, LTD.

By _____
Phillip M. Pippenger, Reg. No. 46055

Leydig, Voit & Mayer, Ltd.
Two Prudential Plaza, Suite 4900
180 North Stetson Avenue
Chicago, Illinois 60601-6780
(312) 616-5600 (telephone)
(312) 616-5700 (facsimile)

RESPONSE UNDER 37 CFR 1.116
EXPEDITED PROCEDURE
EXAMINING GROUP 2121

PATENT
Attorney Docket No. 211163
Client Reference No. PMM2120

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:

Alex Gammerman

Art Unit: 2121

Application No. 09/831,262

Examiner: BELL, MELTIN

Filed: June 29, 2001

For: DATA CLASSIFICATION APPARATUS AND METHOD THEREOF

RESPONSE TO OFFICE ACTION

Mail Stop AF
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In response to the Office Action dated May 26, 2004, please enter the following amendments and consider the following remarks.

CERTIFICATE OF MAILING OR TRANSMISSION UNDER 37 CFR 1.8			
I hereby certify that this Response to Office Action and all accompanying documents are, on the date indicated below, ☒ being deposited with the United States Postal Service as first class mail in an envelope addressed to: Mail Stop AF, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, and also a courtesy copy is ☒ being facsimile transmitted to the U.S. Patent and Trademark Office, Attention: Examiner Bell, Art Unit 2121, Facsimile Number 571-273-3680.			
Name (Print/Type)	Susan Matz		
Signature	Susan Matz	*Date*	November 18, 2004

In re Appln. of:Gammerman et al.
Application No. 09/831,262

AMENDMENTS TO THE SPECIFICATION

Please amend the paragraph on page two starting at line 22 to read as follows:

The present invention provides a new data classification apparatus and method that can cope with high-dimensional classification problems and that provides a universal measure of confidence, valid under the iid (independently and identically distributed) assumption, for each individual classification prediction made by the new data classification apparatus and method.

In re Appln. of:Gammerman et al.
Application No. 09/831,262

AMENDMENTS TO THE CLAIMS

10. A [D]data classification apparatus comprising:

an input device for receiving a plurality of training classified examples and at least one unclassified example;

a memory for storing said classified and unclassified examples;

an output terminal for outputting a predicted classification for said at least one unclassified example; and

a processor for identifying the predicted classification of said at least one unclassified example

wherein the processor includes:

classification allocation means for allocating potential classifications to each said unclassified example and for generating a plurality of classification sets, each said classification set containing said plurality (l) of training classified examples with their classification and said at least one unclassified example ($l+1$) with its said allocated potential classification;

assay means including an example valuation device which determines individual strangeness values (α_i) for each said training classified example ($i=1,2\ldots l$) and said at least one unclassified example ($i=l+1$) having an allocated potential classification (y), the assay means determining a single [an overall] strangeness value ($d(y)$) valid under the [iid] independently and identically distributed assumption for each said classification set in dependence on said individual strangeness values (α_i) of each example by the formula

$$d(y) = \frac{\left|\{i : \alpha_i \geq \alpha_{l+1}\}\right|}{l+1}, \text{ where } i=1,2,\ldots l, l+1;$$

a comparative device for selecting the classification set to which the most likely allocated potential classification for said at least one unclassified example belongs, wherein said predicted classification output by the output terminal is said most likely allocated classification according to said single [overall] strangeness values assigned by said assay means; and

a strength of prediction monitoring device for determining a confidence value for said predicted classification on the basis of said single [overall] strangeness value assigned by said assay means to one of said classification sets to which the second most likely allocated potential classification of said at least one unclassified example belongs.

11. (Canceled)

12. A [D]data classification apparatus as claimed in claim 10, wherein Lagrange multipliers are used to determine said individual strangeness values.

13. (Canceled)

In re Appln. of: Gammerman et al.
Application No. 09/831,262

14. A [D]data classification apparatus comprising:

an input device for receiving a plurality of training classified examples and at least one unclassified example;

a memory for storing said classified and unclassified examples;

stored programs including an example classification program;

an output terminal for outputting a predicted classification for said at least one unclassified example; and

a processor controlled by said stored programs for identifying the predicted classification of said at least one unclassified example, wherein said processor includes:

classification allocation means for allocating potential classifications to each said unclassified example and for generating a plurality of classification sets, each said classification set containing said plurality (l) of training classified examples with their classification and said at least one unclassified example ($l+1$) with its allocated potential classification;

assay means including an example valuation device which determines individual strangeness values (α_i) for each said training classified example ($i=1,2...l$) and said at least one unclassified example ($i=l+1$) having an allocated potential classification (y), the assay means determining a single [an overall] strangeness value ($d(y)$) valid under the [iid] independently and identically distributed assumption for each said classification set in dependence on said individual strangeness values (α_i) of each example by the formula

$$d(y)=\frac{\left|\{i:\alpha_i \geq \alpha_{l+1}\}\right|}{l+1}, \text{ where } i=1,2...l, l+1;$$

a comparative device for selecting the classification set to which the most likely allocated potential classification for said at least one unclassified example belongs, wherein the predicted classification output by said output terminal is the most likely allocated potential classification according to said single [overall] strangeness values assigned by said assay means and

a strength of prediction monitoring device for determining a confidence value for said predicted classification on the basis of said single [overall] strangeness value assigned by said assay means to one of said classification sets to which the second most likely allocated potential classification of said at least one unclassified example belongs.

15. A computer-implemented data classification method comprising:

inputting a plurality of training classified examples and at least one unclassified example;

identifying a predicted classification of said at least one unclassified example which includes,

allocating potential classifications to each said unclassified example;

generating a plurality (l) of classification sets, each said classification set containing said plurality of training classified examples with their classification and said at least one unclassified example ($l+1$) with its allocated potential classification;

determining an individual strangeness value (α_i) for each said training classified example ($i=1,2...l$) and said at least one unclassified example ($i=l+1$) having an allocated potential classification (y), and a single [an overall] strangeness value ($d(y)$) valid under

the independently and identically distributed [iid] assumption for each said classification set in dependence on the individual strangeness values (α_i) of each example by the formula

$$d(y) = \frac{|\{i : \alpha_i \geq \alpha_{l+1}\}|}{l+1}$$ where $i=1,2,\ldots,l,\ l+1$;

selecting the said classification set to which the most likely allocated potential classification for said at least one unclassified example belongs, wherein said predicted classification is the most likely allocated potential classification in dependence on said single [overall] strangeness values;

determining a confidence value for said predicted classification on the basis of the single [overall] strangeness value assigned to one of said classification sets to which the second most likely allocated potential classification for said at least one unclassified example belongs; and

outputting said predicted classification for said at least one unclassified example and said confidence value for said predicted classification.

16. (Canceled)

17. A computer-implemented data classification method as claimed in claim 15, wherein said selected classification set is selected without the application of any general rules determined from the said training set.

18. A data carrier on which is stored a classification program for classifying data by performing the following steps:

generating a plurality of classification sets, each said classification set containing a plurality (l) of training classified examples with their classification and at least one unclassified example ($l+1$) that has been allocated a potential classification;

determining an individual strangeness value (α_i) for each said training classified example ($i=1,2\ldots l$) and said at least one unclassified example ($i=l+1$) having an allocated potential classification (y), and a single [an overall] strangeness value ($d(y)$) valid under the [iid] independently and identically distributed assumption for each said classification set in dependence on said individual strangeness values (α_i) of each example by the formula

$$d(y) = \frac{|\{i : \alpha_i \geq \alpha_{l+1}\}|}{l+1}$$ where $i=1,2,\ldots,l,\ l+1$;

selecting the classification set to which the most likely allocated potential classification for the said at least one unclassified example belongs, wherein the predicted classification is the most likely allocated potential classification in dependence on said single [overall] strangeness values; and

determining a confidence value for said predicted classification on the basis of said single [overall] strangeness value assigned to one of said classification sets to which the second most likely allocated potential classification for said at least one unclassified example belongs.

In re Appln. of:Gammerman et al.
Application No. 09/831,262

REMARKS

Applicant files herewith an amended set of claims, which have been amended to overcome the Examiner's objections.

Claims 10, 12 and 14 have been amended to recite "A data classification apparatus...". Claims 10, 14, 15 and 18 have been amended to ensure that no new matter is added. In these claims, the phrase "an overall strangeness value" has been replaced by the phrase "a single strangeness value", in accordance with page 12, line 10. Claims 15 and 17 now recite "a computer-implemented data classification method" to ensure that they relate to statutory subject matter.

The claims have also been amended to ensure that they are fully distinguished from the prior art, as explained below. In particular, claims 10, 14, 15 and 18 now include the formula by which the single strangeness value is determined by the assay means. This formula is not shown or suggested in any of the prior art documents.

The invention relates to an apparatus for and a method of data classification of unknown items, using a training set of classified examples.

The aim of the invention is not only to predict a classification for an unknown item, but also to provide a measure of confidence in that classification, valid under the iid (independently and identically distributed) assumption. This is the assumption that the training and unknown examples are generated from the same distribution.

As defined in amended claim 10, the data classification apparatus comprises an input device for receiving training classified examples and at least one unclassified example, a memory for storing the examples, an output terminal for outputting a predicted classification for the unclassified example and a processor for identifying the predicted classification, where the processor includes classification allocation means for

allocating potential classifications to each unclassified example and for generating a plurality of classification sets, each set containing the plurality of training classified examples (l) with the classification and at least one unclassified example (l+1) with its allocated potential classification, an assay means including an example valuation device which determines individual strangeness values (α_i) for each training classified example (i=1,2...l) and at least one unclassified example (i=l+1) having an allocated potential classification (y), the assay means determining a single strangeness value (d(y)) valid under the iid assumption for each classification set in dependence on the individual strangeness values (α_i) of each example by a given formula, a comparative device for selecting the classification set to which the most likely potential classification for the unclassified example belongs on the basis of the single strangeness value assigned by the assay means, and a strength of prediction monitoring device for determining a confidence value for the predicted classification on the basis of the single strangeness value assigned to the classification set to which the second most likely potential classification of the unclassified example belongs.

The formula $d(y) = \frac{\left|\{i : \alpha_i \geq \alpha_{l+1}\}\right|}{l+1}$, where i=1,2....l, l+1, calculates the single strangeness value by counting how many of the individual strangeness values of the training examples are greater than or equal to the individual strangeness value of the unclassified example, and dividing this by the total number of examples.

Thus, the apparatus starts with training classified examples, and uses these to predict the classification of an unclassified example. The classification allocation mean allocates potential classifications to each unclassified example, and generates a number of classification sets, each containing the said training examples with their classification, and at least one unclassified example with a potential classification. Each pair of the example and classification in the set is tested by the assay means to determine its individual strangeness value. That is a measure of how strange it would be if the classification were correct. The assay means then determines a single strangeness value,

valid under the iid assumption, for each classification set, in dependence on the individual strangeness values. The comparative device compares the single strangeness value for all the sets, to see which is most likely to be correct for the unclassified example. The strength of prediction monitoring device then looks at the second most likely classification, on the basis of the single strangeness values, to give a measure of confidence that the most likely classification is in fact correct. Thus, the greater the difference between the single strangeness values of the most likely potential classification and the second most likely potential classification, the greater is the measure of confidence that the most likely classification is correct.

We submit that this combination of features, and in particular the formula, is not shown in any of the prior art documents.

Applicant agrees that Mizuno shows an input device, a memory, an output terminal and a processor, with the processor including classification allocation means and assay means. However, applicant submits that Mizuno calculates an error between a classification assigned to an unknown pattern and a correct classification, and initiates re-training when the average error exceeds a predetermined value. This copes with the possibility that the unknown patterns are generated by a different mechanism from the training patterns (see column 1, lines 9 to 19). This differs from the invention as claimed, which specifies that the single strangeness value is valid under the iid assumption, in that the unknown and training examples are generated from the same distribution. It is submitted that column 7, lines 1 to 12 of Mizuno do not show the determination of an individual strangeness value for each training example and an unknown example and then the determination of a single strangeness value for each classification set, in dependence on the individual strangeness values. In particular, Mizuno does not show the formula now included in amended claim 10. Applicant agrees that Mizuno does not teach a strength of prediction monitoring device for determining a confidence value for the most likely classification, on the basis of the single strangeness value of the second most likely potential classification.

It is submitted therefore that claim 10 is clearly novel over Mizuno. Further, Mizuno does not show or suggest determining a single strangeness value valid under the iid assumption, and indeed is concerned with the possibility that the unknown patterns are generated from a different distribution, rather than the same distribution. Thus, Mizuno teaches away from the invention, rather than towards it. In addition, Mizuno does not show or suggest determining individual strangeness values and then a single strangeness value, either in general or on the basis of the claimed formula, nor the strength of prediction monitoring device for a confidence value on the basis of the single strangeness value of the second most likely potential classification.

Tsuboka relates to pattern recognition employing the Hidden Markov Model, and is particularly relevant to speech recognition or other time series. In such pattern recognition it is necessary to identify a spoken word, and to decide which of a number of classified words it is most likely to be. Examples of classified words are put into clusters and then the probability is calculated that the word belongs to a particular cluster. Tsuboka recognises that a pattern to be recognised has different probabilities of belonging to different clusters, and calculates the probability in various ways, using a comparative device to select the most likely classification. However, the aim of Tsuboka is simply to find the cluster with the highest probability, for accuracy of recognition.

Applicant submits that firstly Tsuboka does not use the iid assumption, as claimed by the applicant. Secondly, applicant submits that Tsuboka does not use the second most likely classification as a strength of prediction monitoring device. The passage indicated by the Examiner (column 11 lines 11 to 20) appears simply to say that a pattern could be in more than one classification with different probabilities. The subsequent portion of the specification indicates how the most likely classification is then computed. Tsuboka clearly does not show the use of the formula, as now included in amended claim 10.

Thus, it is submitted that claim 10 is clearly novel over Tsuboka. Further, Tsuboka does not show or suggest using the second most likely allocated potential

classification as a strength of prediction monitoring device for determining a confidence value of the predicted classification. Claim 10 is therefore clearly inventive in view of Mizuno.

With regards to a combination of Mizuno and Tsuboka, it is submitted that they would not lead to the present invention. Firstly, neither is concerned with ensuring that the strangeness value is valid under the iid assumption, and indeed Mizuno is concerned with the possibility that the unknown patterns are generated from a different distribution. In addition, neither Mizuno nor Tsuboka shows the use of the second most likely classification as the strength of prediction monitoring device to provide a confidence value for the correctness of the predicted classification.

Thus, a combination of Mizuno and Tsuboka, even though they are both concerned with maintaining accuracy of the classification, simply would not lead to the present invention, as claimed in claim 10, because of the lack of provision of a confidence value using the second most likely classification, and the fact that the strangeness values are not valid under the iid assumption. The combination also fails to show the use of the formula enumerated in the claim.

Applicant submits therefore that claim 10 is clearly distinguished from all the prior art, and is therefore allowable. Claims 14, 15 and 18 have the same limitations as claim 10, and it is therefore submitted that these are also allowable.

As claims 12 and 17 are dependent on claims 10 and 15, it is submitted that these are allowable to the same extent.

In re Appln. of:Gammerman et al.
Application No. 09/831,262

Conclusion

The application is considered in good and proper form for allowance, and the Examiner is respectfully requested to pass this application to issue. If, in the opinion of the Examiner, a telephone conference would expedite the prosecution of the subject application, the Examiner is invited to call the undersigned attorney.

Respectfully submitted,

Phillip M. Pippenger, Reg. No. 46,055
LEYDIG, VOIT & MAYER, LTD.
Two Prudential Plaza, Suite 4900
180 North Stetson Avenue
Chicago, Illinois 60601-6780
(312) 616-5600 (telephone)
(312) 616-5700 (facsimile)

Date: November 18, 2004

Amendment or ROA - Final (Revised 10/21/2004)